Minutes of IARC meeting 99, May 9th, 2022

In attendance: Ayelet Peres, Gur Yaari, Andrew Collins, Martin Corcoran, William Lees, Corey Watson, Mats Ohlin, James Heather (guest)

1. **Approval of minutes of meeting 97**
   Approved

2. **Approval of minutes of meeting 98**
   Approved

3. **AIRR-C meeting**
   Discussion of presentation of IARC-SC actions

4. **Functional groups reference book - update**
   Brief presentation of current updates by AP

5. **Approach to inference of the 3'-end**
   IARC intends to approve sequences as far towards the 3'-end as we can be confident and to highlight the possible downstream bases that are likely there. The opportunity to subsequently provide two reference sets, one with affirmed sequences (even if short) and one with most likely full length sequences for use in tools that require full length sequences for optimal functionality.

   Discussion on practical outcome of inference of 3'-end inference as exemplified by data supporting inference of TRBV19*01_A24G, TRBV12-4*01_C87T, TRBV5−6*01_T284G, TRBV7-7*01_C315T, TRBV12-5*01_C28G_T140A, TRBV6-6*01_G264A, and TRBV15*02_G275A, a difficult process in the case of inference of immunoglobulin light chain and TCR sequences in which substantial 3'-end processing occurs. It was decided that we, as we do for IGHV alleles, report likely bases 3' of the last solidly inferred allele.

6. **Assessment of inference TRBV19*01_A24G in P4_I24_S1 (S00036)**

   TRBV19*01_A24G has been inferred in six genotypes in the VDJbase P4 data set, including in VDJbase P4_I24_S1, a haplotypable data set (based on heterozygocity in TRBJ1-6). The genotype is also implied to carry

TRBV19*01. No other gene in the IMGT database is closely related to these alleles of TRBV19. The novel allele is the most expressed allele in the repertoire (53% allelic frequency; 0.71% of the total error-free population). It is represented by 240 error-free sequences and 202 unique CDR3s in the error-free set. Haplotyping based on allelic diversity in TRBJ1-6 demonstrates perfect separation from TRBV19*01. IARC affirms the sequence at Level 1 up to and including base 322. It is acknowledged that the allele most likely carries 4 additional bases, typically TAGA, at base positions 323-326. Trailing "." indicates IARC's opinion that the sequence is likely to contain additional 3'-nucleotides for which there is insufficient evidence to make an affirmation.

```
>TRBV19*01_A24G
GATGGTGGAATCACTCAGTCCCCGAAGTACCTGTTCAGAAAGGAAGGACAGAATGTG
ACCCTGAGTTGTGAACAGAATTTGAACCACGATGCCATGTACTGGTACCGACAGGAC
CCAGGGCAAGGGCTGAGATTGATCTACTACTCACAGATAGTAAATGACTTTCAGAAA
GGAGATATAGCTGAAGGGTACAGCGTCTCTCGGGAGAAGAAGGAATCCTTTCCTCTC
ACTGTGACATCGGCCCAAAAGAACCCGACAGCTTTCTATCTCTGTGCCAGTAGTA..
..
```

Result summary: TRBV19*01_A24G	No rearrangement found		
V-GENE and allele	Homsap TRBV19*01 F	score = 1356	identity = 99.63% (272/273 nt)
FR-IMGT lengths, CDR-IMGT lengths	[5.6.X]		

1. Alignment for V-GENE and allele identification

Closest V-REGIONs (evaluated from the V-REGION first nucleotide to the 2nd-CYS codon)

```
                              Score    Identity
L36092   Homsap TRBV19*01 F    1356    99.63% (272/273 nt)
M97725   Homsap TRBV19*03 (F)  1347    99.27% (271/273 nt)
U48259   Homsap TRBV19*02 F    1338    98.90% (270/273 nt)
L33102   Homsap TRBV10-3*04 [F] 510    65.20% (178/273 nt)
U08314   Homsap TRBV28*01 F     510    65.20% (178/273 nt)
```
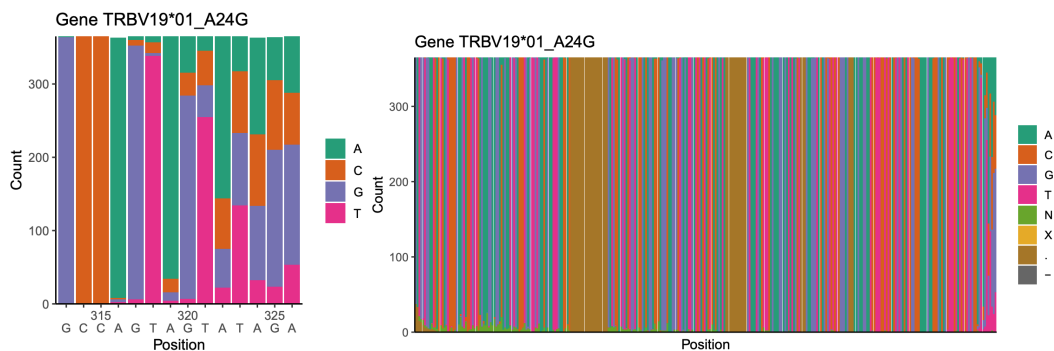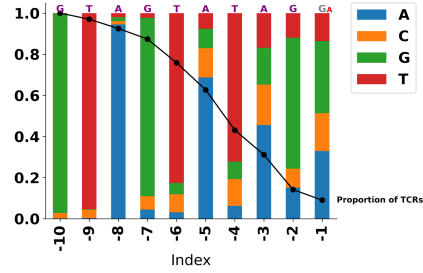

Consensus plot of 3'-end of alleles of TRBV19:
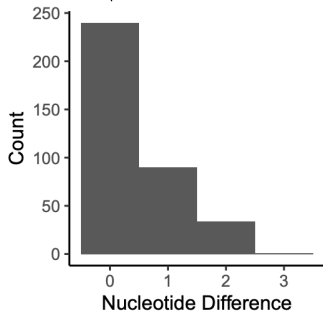
All subjects, all alleles
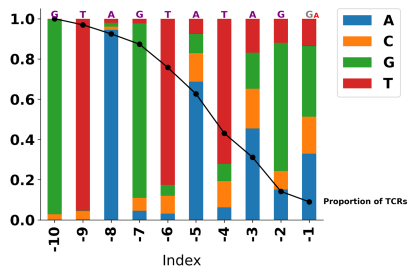
One subject, TRBV19*01_A24G
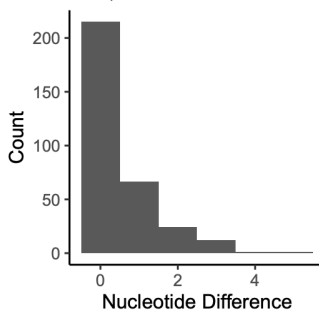
TRBV19*01_A24G

365 sequences assigned
240 (65.8%) exact matches, in which:
202 unique CDR3
13 unique J

TRBV19*01
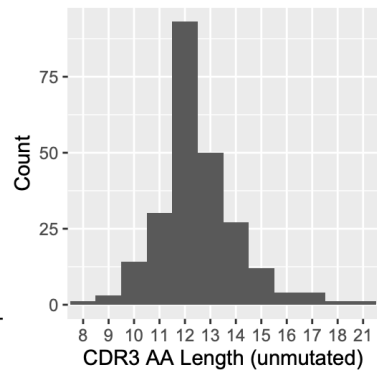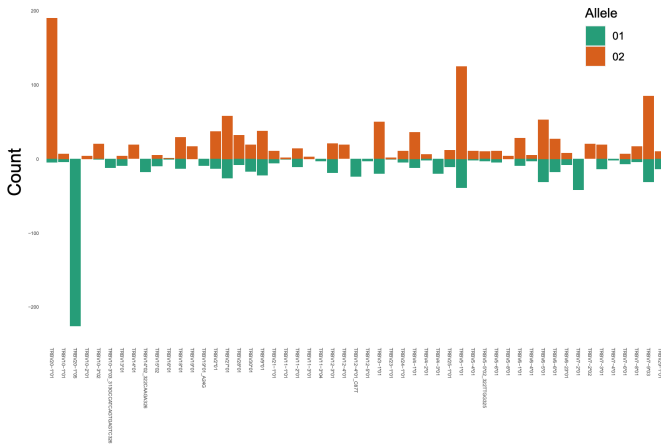
319 sequences assigned
215 (67.4%) exact matches, in which:
162 unique CDR3
13 unique J

TRBV19*01_A24G

Sequence Count by TRBJ1–6 allele usage

7. **Date for next meeting is to be decided after the AIRR-C meeting**